

## Durham Research Online

---

### Deposited in DRO:

10 February 2015

### Version of attached file:

Accepted Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Beckmann, N. and Beckmann, J.F. and Birney, D. P. and Wood, R. E. (2015) 'A problem shared is learning doubled : deliberative processing in dyads improves learning in complex dynamic decision-making tasks.', Computers in human behaviour., 48 . pp. 654-662.

### Further information on publisher's website:

<http://dx.doi.org/10.1016/j.chb.2015.01.052>

### Publisher's copyright statement:

NOTICE: this is the author's version of a work that was accepted for publication in Computers in Human Behavior. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Computers in Human Behavior, 48, July 2015, 10.1016/j.chb.2015.01.052.

### Additional information:

## Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

A Problem Shared is Learning Doubled: Deliberative Processing in Dyads Improves  
Learning in Complex Dynamic Decision Making Tasks

Nadin Beckmann<sup>a</sup>, Jens F. Beckmann<sup>a</sup>, Damian P. Birney<sup>b</sup>, Robert E. Wood<sup>c</sup>

*Corresponding author:*

Dr Nadin Beckmann

<sup>a</sup> Durham University

Durham DH1 1TA, UK

<sup>b</sup> University of Sydney

Sydney, NSW 2006, Australia

<sup>c</sup> University of Melbourne

Melbourne, VIC 3053, Australia

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting and proofreading process, which might lead to differences between this version and the version of record.

Beckmann, N., Beckmann, J.F., Birney, D.P., & Wood, R.E. (in press). A problem shared is learning doubled: Deliberative processing in dyads improves learning in complex dynamic decision-making tasks. *Computers in Human Behaviour*.

## Abstract

Whilst micro-worlds or simulations have increasingly been used in higher education settings, students do not always benefit as expected from these learning opportunities. By using an experimental-control group design we tested the effectiveness of structuring the task environment so as to encourage learners to approach simulations more systematically. Seventy-one professionals who participated in a postgraduate-level management program worked on a management simulation either individually ( $n = 35$ ) or in dyads ( $n = 36$ ) while exploring the simulation (exploration phase). Peer interactions in the shared learning condition were structured so that learners were encouraged to employ hypothesis-testing strategies. All participants then completed the simulation again individually so as to demonstrate what they had learned (performance phase). Baseline measures of cognitive ability and personality were also collected. Learners who explored the simulation in the shared learning condition outperformed their counterparts who explored the simulation individually. A simple manipulation of the way learners interacted with the simulation facilitated learning. Improved deliberation is discussed as a potential cause of this effect, preliminary evidence is provided. This study lends further evidence that the effectiveness of learning using simulations is co-determined by characteristics of the learning environment.

*Keywords:* simulations, hypothesis testing, adult learning, learning environment, learner dyads

# A Problem Shared is Learning Doubled: Deliberative Processing in Dyads Improves Learning in Complex Dynamic Decision-Making Tasks

## 1 Introduction

Micro-world simulations have been utilised to date in various higher education settings, for example in medical and nursing education (McGaghie, Issenberg, Petrusa, & Scalese, 2006; Ravert, 2002; Karakus, Duran, Yavuz, Altintop & Caliskan, 2014), business and management education (Lainema & Nurmi, 2006; Pasin & Giroux, 2011; Beckmann, Minbashian, Wood, & Tabernero, 2012; Romme, 2003; Zantow & Knowlton, 2005), engineering (Chung, Harmon & Baker, 2001; Fang, Tan, Thwin, Tan, & Koh, 2011; Mendonca, Chang, Hu, & Gu, 2012), and more recently in social work education (Wastell, Peckover, White, Broadhurst, Hall, & Pithouse, 2011). The purpose of using simulations in teaching and learning varies with regard to the learning objectives, ranging from the acquisition of domain specific skills (e.g., flight simulators) to the acquisition of domain general skills of systematic enquiry, such as hypothesis testing. However, perhaps surprisingly, students often underperform in these learning environments and as a consequence, do not always benefit as expected from the use of simulations in their learning (e.g., Groessler, 2004). Potential causes for this phenomenon can be conceptualized in terms of three dimensions: the learner, the simulation, and the situation. Individual differences in learners' prior knowledge, levels of expertise, motivation, and reasoning ability and their relationships to learning outcomes have primarily been the focus of psychological research. Features of the simulated micro-worlds such as delays, feedback loops, and non-linearities are

heavily featured in research in system dynamics; whilst attributes of the learning environment (e.g., how information is presented) are primarily addressed in research with an instructional design focus. In our study we explore how prescribing the way individuals interact with a simulation affects learning behavior and subsequent learning outcomes.

### *1.1 Micro-worlds as Learning Tools*

Micro-worlds are task environments attempting to simulate (more or less) comprehensively real-world problems and their underlying principles. Typically these are complex, open-ended problems that require learners to make decisions, monitor the outcomes of their decisions, and learn from feedback. As an example, the Furniture Factory Simulation (Goodman & Wood, 2004) is a computer-based environment that simulates the motivational processes at play in managing a group of employees over several simulated weeks of business activity. The learning objective is to gain an understanding of the interplay between managerial decisions and various motivational responses by virtual employees. It is expected that the decisions learners make when working on simulation tasks are the types of decisions professionals would make on the job.

Simulations vary widely in their domain and task characteristics. Regardless of these differences, researchers typically purport several benefits of using simulations in training and education (e.g., Wood, Beckmann, & Birney, 2009). Amongst those are: (1) Simulations represent a safe learning environment in which the impact of decisions is modelled, but obviously not realized. This provides the opportunity to experiment with different decision strategies in a risk free environment where there are no costs associated with potentially poor decision-making. (2) The use of

simulations is expected to be engaging and motivating, because simulations promise  
 a meaningful approximation to authentic problem solving (e.g., Chang, Peng &  
 Chao, 2010). (3) Simulations are expected to enable learners to link theory and  
 practice. Learning with simulations seems to promise an experiential  
 contextualisation of ‘textbook knowledge’. (4) The use of simulations is thought to  
 foster self-directed learning. Learning is self-directed in situations where the learner  
 (rather than a tutor) is in control of the learning experience (Gureckis & Markan,  
 2012). For example, a student that actively searches for information that is not readily  
 available engages in self-directed learning. When students work on simulations they,  
 to some extent, determine which information they are exposed to depending on the  
 decisions they make. In this regard, simulations also represent a snapshot of the real  
 world where employees are often expected to continue learning on the job with  
 minimal guidance. (5) Simulations are also believed to help students to practice  
 important cognitive and meta-cognitive skills that are involved in successful  
 problem-solving, such as systematic hypothesis testing and exploration (Beckmann  
 & Goode, 2014; Burns & Vollmeyer, 2002). In sum, the use of simulations in higher  
 education contexts is expected to engage and motivate students, to encourage  
 students to contextualise their knowledge, and to practice problem-solving skills that  
 are applicable across a wide range of contexts.

In contrast to the many suggested benefits of using simulations in learning, the  
 evidence as to whether students actually learn effectively when working on these  
 tasks is mixed (Bell, Kanar, & Kozlowski, 2008; Gosen & Washbush, 2004; for an early  
 review see Lane, 1995). Some studies have found simulations to provide effective  
 learning environments (Chung et al., 2001; Ravert, 2002), others were unable to

replicate such findings (Gresse van Wangeheim, Thiry, & Kochanski, 2009; Stouten, Heene, Gellynck & Polet, 2012; Njoo & DeJong, 1993; Qudrat-Ullah & Karakul, 2007, see also the discussion on poor performance of participants in problem-based learning environments in general, Ellis, Marcus & Taylor, 2005). Groessler (2004) identified no less than 15 issues concerning the use of simulations as teaching and research tools, including research design and methodological obstacles to evaluating the effectiveness of simulations as learning tools, task difficulty due to complexity, and, depending on individual differences in cognitive ability and knowledge, difficulties students often have in making sense of the task. The latter can lead to random decision-making, which impedes any learning.

It is clear that an evaluation of the evidence for or against the effectiveness of simulations as learning tools needs to reflect on various challenges. These challenges include methodological constraints as well as conceptual shortcomings. With regard to the former, one major issue is that quite a few studies that report positive effects on learning do not employ study designs that would allow such conclusions. Many studies, for instance, lack a control group (e.g., Adobor & Daneshfar, 2006; Cronan et al., 2011; Chung, et al., 2001; Hung, 2008; Qudrat-Ullah, 2010), which challenges the validity of claims that reported performance increases can in fact be attributed to the use of the particular simulation. Another challenge is ambiguity in what constitutes an indicator of learning success. Studies variously report on self-perception of learning, knowledge tests, causal diagrams, various performance indicators within the simulation, and performance in transfer problems or so-called real-life outcomes.

Student motivation, and as a consequence student engagement, are often reported to be high when simulations are employed (e.g., Shellman & Turan, 2006;

Chang, et al., 2010). However, this does not necessarily translate into better learning (Stouten et al., 2012; Adobor & Daneshfar, 2006). For instance, Stouten et al. report that whilst learners had confidence in the simulation, found it a valid model of reality, and believed that they had learned important content, no learning was observed with regard to objective learning outcomes (e.g., changes in participants' knowledge). Indeed, students often perform relatively poorly in simulations (Paich & Sterman, 1993). Also performance indicators derived from within the simulation are not necessarily a valid indicator of learning success. "Game performance" scores often reflect success in pursuing some sort of optimisation goal (e.g., maximising market share or minimising staff costs). Achieving good performance scores requires decision-making behavior which is different from exploration behavior geared towards the acquisition of structural or functional knowledge about the relatedness of decision variables and outcome variables in a simulation. In other words, the operationalization of game performance tends to reward a different kind of behavior than what these scores are supposed to be indicative of (i.e., learning behavior).

Transfer scores (i.e., performance success outside the simulation itself) can be seen as indirect learning indicators at best because success in learning within a simulation does not always translate into success in other tasks or 'real-world' complex, dynamic problems. Beckmann and Goode (2014) have proposed that such lack of transfer might be caused by one of the core features of simulations, namely their attempt to provide a contextualised learning environment by using semantically meaningful variable labels and cover stories. It can be argued that the emphasis on contextualisation of learning with simulations comes with the risk that learning outcomes (i.e., knowledge and understanding) achieved in more or less



narrowly defined contexts are less likely to be utilised in novel, albeit homomorphous real-life situations (Beckmann & Goode).

Various reasons are discussed for the limited effectiveness of simulations as learning tools. One potential reason for the ‘under-performance’ of students in simulations is that students are cognitively overwhelmed by the complexity of the task (Gonzalez, 2005; Gonzalez, Vanyukov & Martin, 2005; Wood, et al., 2009). Several studies have tried to address this issue (e.g., Lerch & Harter, 2001). Parush, Hamm and Shtub (2002) provided learners with a ‘learning history tool’ which allowed access to and tracking of their past decisions and subsequent effects. Externalisation of the decision history was discussed as one way to reduce cognitive demands of the task which led to better performance compared to a control group who played the simulation in the traditional way.

Another factor that contributes to an under-utilisation of learning opportunities in simulations is the tendency of learners’ to inadequately encode the complexity of the task by focussing on surface features of the simulation and by producing high decision-making densities, i.e., acting too quickly and therefore unreflectively. Such passive encoding of simple action-outcome relationships is counterproductive to the main learning objective when working with simulations, namely the acquisition, the elaboration, and the refinement of cognitive schemata. The acquisition of such schemata or adequate mental models that underpin successful decision-making in complex environments (Goode & Beckmann, 2010) requires sustained deliberative processing of information. Deliberative processing requires the use of working memory resources to store and recall the outcomes of previous actions (Anderson & Schunn, 2000) including making explicit predictions and comparing them with actual

outcomes. As argued elsewhere (Beckmann & Goode, 2014, p. 287), the effectiveness of (discovery) learning using simulations (e.g., de Jong & van Joolingen, 1998) is challenged by a number of factors. These include difficulties learners seem to have in hypothesis generation, in designing of experiments, and in interpreting decision outcomes (de Jong, 2006; Njoo & de Jong, 1993; for an early analysis see Wolfe, 1975). Pedagogical efforts to scaffold necessary systematicity in decision-making behavior include the implementation of a “hypothesis scratchpad” (van Joolingen & de Jong, 1993) that is meant to facilitate the generation of hypotheses. However, even the provision of quite explicit hints as to what kind of decision sequences are likely to generate most informative system states (Leutner, 1993; Jansson, 1995) have not resulted in a convincingly consistent effect pattern (de Jong & van Joolingen, 1998, p. 193).

Evidently, just “playing the simulation” does not necessarily lead to learning (e.g., Gresse van Wangenheim et al., 2009; Pfahl et al., 2001, Leemkuil & de Jong, 2012). “Micro-worlds do not work on their own, meaning that there probably needs to be some structuring of the participants’ interactions with the micro-world to obtain (or increase) learning effects” (Stouten, et al. 2012, p. 768, see also Bell et al., 2008). The aim of this study was to contribute to our understanding of how to embed guidance and support structures into simulation-based learning. We investigate whether it is possible to help learners better exploit learning opportunities by (a) approaching and exploring the learning task in systematic ways, and (b) engaging in deliberative processing of information.

## 1.2 Our Approach

The objective was to encourage deliberative processing and to prevent hasty, unreflective actions that potentially impede learning in terms of the acquisition of adequate mental models for decision-making processes. Drawing on the purported benefits of peer interaction for learning (e.g., Lou, Abramson, & Apollonia, 2001), we designed a learning setting where learners work in dyads, and where learners were instructed to take on either of two roles at a given time, the role of the ‘decision-maker’ or the role of the ‘decision-executer’. The task of the ‘decision-maker’ was to decide on the next input (i.e., the next set of interventions); the task of the ‘decision-executer’ (operator) was to enter the decision into the computer and to prompt the decision-maker to (a) explain and justify their decision and to (b) explicitly state the expected outcome. This approach serves two main purposes. First, it prevents learners from acting prematurely before considering potential consequences of their decisions (i.e., from using a ‘trial and error’ strategy). Second, the necessity to communicate plans and decisions with the peer learner facilitates conscious deliberation of decisions and their consequences.

The aim of this study was to test our assumption that learners in the dyad-setting better utilise learning opportunities of the simulation than learners in the individual setting (control). By comparing performance scores and decision times between the dyad group and the control group, we test for effects on three levels. We tested for a *proximal learning effect*, a *distal learning effect*, and for a *deliberation effect*. A proximal learning effect refers to performance improvements within a simulation run (i.e., across the 16 decision cycles or trials that constitute a run). A distal learning effect refers to performance improvements across 4 simulation runs (each consisting

of 16 trials). The deliberation effect refers to the amount of time learners invest in their decision-making across runs.

## 2 Material and Methods

### 2.1 Participants

We recruited a sample of 71 professionals who participated in a postgraduate-level management training and development program. The sample is a subset drawn from the *A.L.L. Flexible Expertise data base*<sup>1</sup> ( $N = 423$ ). Participants were working in middle-level management roles (aged 25 to 49 years,  $M = 35.53$ ,  $SD = 6.58$ , 43% female) at one of three large Australian companies (an international bank, a major international airline, a national broadcasting company). On average participants had about six years of experience in management and had worked about two years in their current role within the respective organisation. Of these, 70% had completed a university degree (33% postgraduate; 37% undergraduate) and 17% reported “high school” as their highest level of education. The remaining 13% of participants reported having completed a different degree (“other”).

### 2.2 Measures

#### 2.2.1 Simulation

The Employee Management Simulation (EMS) is a computer-based environment that simulates the process of managing the performance of an employee over 16 simulated weeks of business activity. The EMS is based on the Furniture Factory Simulation (Goodman & Wood, 2004; Wood & Bailey, 1985).

The EMS includes three decision variables (i.e., goal, guidance, and reward) and two outcome variables (i.e., employee motivation and employee job performance). Participants are required to assign performance targets (i.e., “goal”) and to provide guidance and rewards to motivate their employee, and then monitor the impact of these decisions on the employee’s simulated performance. Figure 1 provides a screenshot of the simulation. Participants enter their decisions for each simulated week of business activity. A simulated week of business activity is referred to as a “trial”. The outcome variables are then calculated based on the underlying algorithm of the simulation. The underlying algorithm was a system dynamics model that reflects the motivational processes research has indicated as general principles underlying employee performances (e.g., expectancy-value theory). The past and current outcomes of each decision trial are displayed on the computer screen to allow participants to learn from their decision history. The objective of the EMS is to make decisions that motivate the simulated employee to perform well. The outcome variable *employee job performance* – operationalized as job completion time (reverse scored) – is interpreted as an indicator of simulation-based learning and serves as the core dependent variable in the current study. This variable will be referred to as *performance* in the remainder of the paper. We also recorded the time it took participants to make their decisions (decision time) for each simulated week of business activity (i.e., trial).

### 2.2.2 *Individual Differences Measures*

As part of the management expertise training and development program, participants were assessed on a range of cognitive and personality variables. Details of these tests as well as reliability estimates are provided below.

*Cognitive ability* was assessed using SHL verbal, numerical and abstract reasoning tests. *Verbal Reasoning* (SHL-VR) is a 48-item test that measures the ability to understand and evaluate the logic of various verbal arguments, relevant to managerial work. The task was to decide whether a statement made in connection with the given information was true, untrue, or whether there was insufficient information to judge. The score analysed is the number of items answered correctly (Cronbach's  $\alpha = .82$ ). *Numerical Reasoning* (SHL-NR) is a 35-item test that measures the ability to make inferences based on numerical data and was designed to apply to a range of management level jobs. The task was to interpret data and combine information from different sources in order to answer the questions given. Calculators were provided so that the emphasis in this test was on understanding and evaluation rather than on computation. The score analysed is the number of items answered correctly (Cronbach's  $\alpha = .91$ ). *Abstract Reasoning* (SHL-AR) is a 40-item test that measures the ability to reason with abstract figures and requires the recognition and application of logical rules governing sequence changes. The abstract reasoning test consisted of a series of diagrammatic sequences. The task was to identify the underlying structure of this sequence, and select the figure that best completed the pattern. The score analysed is the number of items answered correctly (Cronbach's  $\alpha = .85$ ).

*Personality* was assessed using the International Personality Item Pool (IPIP) version of the NEO inventory (Goldberg et al., 2006; see [http:// ipip.ori.org/](http://ipip.ori.org/) ). The IPIP NEO inventory is based on the five-factor model of personality (Costa & McCrae, 1992) and contains 50 items assessing five broad dimensions of personality (neuroticism, conscientiousness, agreeableness, openness to experience and

extraversion). Participants were instructed to describe themselves as they generally are compared to other people of the same sex and roughly the same age. The answer format for all items was a visual analogue scale that required participants to place a marker along a line with the polar ends labelled “strongly disagree” to “strongly agree” (Cronbach’s  $\alpha_N = .85$ ,  $\alpha_C = .87$ ,  $\alpha_A = .76$ ,  $\alpha_O = .78$ ,  $\alpha_E = .88$ , as assessed in a larger sample of which the current participants are a sub-sample,  $N = 423$ , see footnote 1).

### 2.3 Design

An experimental-control group design was employed with *condition* (shared learning versus individual learning) as a between-person factor, and simulation run (1 to 4) as a within-person factor.

The study was carried out in a context analogous to higher education where individuals who participated in the same teaching session needed to complete the task under the same experimental condition and therefore cluster sampling was used. Learner cohorts (i.e., clusters) were randomly assigned to either the experimental condition ( $n_{shared} = 36$ ) or control condition ( $n_{individual} = 35$ ). Gender ratio and age were comparable across both groups. One participant from the shared learning group withdrew from the programme and this individual’s data is not included in the analyses.

#### 2.3.1 Experimental factor: Shared vs individual learning manipulation

The design was introduced in an attempt to improve learning using simulations with an emphasis on deliberative processing. During the exploration phase, participants were instructed to complete two simulation runs. In the *shared*

*learning condition*, participants worked in dyads. In simulation run 1, participant A's task was to decide on the actions to be taken at each trial ('decision-maker' role). Participant B was instructed to (a) elicit a justification from participant A for their decisions (i.e., explanation and prediction of possible consequences), (b) provide feedback on these, (c) enter the decision into the computer ('executer' role), and finally (d) to feed back the simulated decision outcomes. The roles of 'decision-maker' and 'executer' were swapped for run 2 of the simulation. This approach was chosen to create a condition where participants were encouraged to deliberately and explicitly elaborate on the causes and consequences of their decisions while working on the simulation (i.e., hypothesis testing). In other words, the aim was to prevent participants from acting hastily and making decisions prematurely without having considered the potential consequences of their actions (i.e., to encourage learners to deliberately process information and to act reflectively). In the *individual learning condition* learners worked on the simulation individually for two runs; this is in line with a 'learning-by-doing' strategy. All learners (experimental as well as control group) then worked on the simulation *individually* for another two simulation runs. These two simulation runs constitute the *performance phase* of the experiment.

## 2.4 Procedure

The study comprised two sessions. In Session 1, learners completed the ability tests and personality measures. Learner cohorts were then randomly allocated to either the shared learning condition or the individual learning condition. The learning task using the simulation with its exploration phase and performance phase was completed in Session 2.



At the beginning of Session 2, all learners (experimental and control group) were instructed to employ a ‘hypothesis testing’ strategy and were informed about the benefits of such a strategy when working on complex, dynamic problems. As part of this pre-briefing, learners were also instructed on the usefulness of making explicit predictions regarding the anticipated outcomes of their decisions. Their attention was drawn particularly to identifying any discrepancies between predictions and actual outcomes because they represent feedback that can be used (a) to modify/ fine-tune their mental models under development (i.e., knowledge acquisition and learning) and (b) to inform their subsequent decisions in terms of an effective exploration strategy. In the shared learning condition learners were also instructed on the distinct roles they were expected to adopt during the exploration phase. All learners were introduced to the EMS via an on-line tutorial and then continued on to work either individually (individual learning condition) or in dyads (shared learning condition) for two runs (16 trials each). Whilst the instructional focus for the two runs in the exploration phase was on learning, the instructional focus for the subsequent two runs was on performance. For these two runs in the performance phase all learners worked on the simulation individually. This procedure ensured that (a) the exposure time to the simulation was the same for both the dyad learning group and the individual learning group during the exploration phase (i.e. two runs for every learner) and (b) that performance scores were obtained under individual work conditions for every learner. No time pressure was put on learners in either condition. On average participants took about 90 min to complete all simulation-related tasks. Learners in both conditions knew each other from a

previous teaching session they attended six months prior to this session. The learner pairings were created ad hoc for this learning exercise.

Performance scores in the first two (exploration) runs were averaged for learners in the individual learning condition, performance scores for dyad-learners were based on the run where the individual learner served as the decision-maker. Thus, in effect, data for three runs of 16 trials (simulated weeks) per learner were available for analysis; a combined exploration run and two performance runs.

## 2.5 Data Analysis

To account for the nested structure of the data, hierarchical linear modelling (HLM, Raudenbush & Bryk, 2002) of the variation in the performance outcome variable was carried out. Trials (16 weeks of simulated business activity per simulation run) were nested within simulation runs, which were nested within individuals ( $N = 70$ ). First, an unconditional model (Model 1) was calculated to assess the variance associated with each level in the data structure.

To test for the *proximal learning effect* as well as the *distal learning effect*, a three-level hierarchical linear model (Model 2) was estimated in which the dependent variable was performance and the independent variables were trial (i.e., simulated week of business activity) on level 1, simulation run (1+2 to 4) on level 2, and condition (shared or individual learning in the exploration phase, i.e. run 1+2) on level 3. Level-1 and level-2 independent variables were group-mean centred, the level-3 independent variable was grand-mean centred. To further specify the effect of the experimental condition on performance the following cross-level interaction effects were also estimated: condition x trial and condition x run. The *condition x trial* effect tests whether those in the shared learning condition showed superior learning

as indicated by a steeper increase in performance as they work through a simulation run (trial 1 to trial 16), compared to those in the individual learning condition (*proximal learning* effect). The *condition x run* effect allows testing whether those in the shared learning condition benefited more from the learning experience as indicated by a higher performance increase from simulation run 1+2 to run 4, compared to those in the individual learning condition (*distal learning* effect). Robust standard errors were used in analysing the significance of the main and interaction effects. Reliability estimates of level 1 and level 2 coefficients were checked and deemed appropriate (ranging from .65 to .94).

In order to explore whether the experimental manipulation influenced learners' exploration behavior (i.e., *deliberation effect*), we also analysed their decision times as a proxy for their amount of reflectiveness. An analogous model (see Model 2) was estimated, however, with decision time as the dependent variable (Model 3). Response times were windsorised (criterion: 3 SDs) in order to ensure that large outliers would not unduly influence the findings. Reliability estimates of level 1 and level 2 coefficients were checked and deemed appropriate (ranging from .52 to .71). We present the equations that describe models 2 and 3 in Appendix A.

The adopted analysis strategy focuses on trends in performance scores over time. This goes beyond the conventional comparison of final game scores and follows Bell et al.'s (2008, p. 1429) plea for employing more process-oriented approaches in our efforts to examining the impact of different features of simulations on learning.

### 3 Results

Table 1 shows the means, standard deviations and correlations for all study variables at the between person level (HLM level 3). We first tested whether there were any systematic differences in cognitive ability and personality variables between the experimental and the control group. No such differences were observed.

-----

insert Table 1 here

-----

Using HLM we then estimated the proportion of variance observed in performance at each level in the data structure (Model 1, unconditional model). 30.33% of the total variance observed occurred at level 1 (trials 1 to 16) and 54.21% at level 2 (simulation run 1+2 to run 4). In other words, the major proportion of the variance observed in performance (84.54%) occurred within individuals ( $\text{var}_{\text{total}} = 335.47$ ,  $\text{var}_{\text{within}} = 283.63$ ) rather than between individuals ( $\text{var}_{\text{between}} = 51.84$ ). These findings suggest that the performance of an individual varies more during task completion (from trial to trial, and from simulation run to simulation run) than performances of different individuals vary from another.

We then estimated the main and interaction effects of condition and simulation run on performance in a given trial (Model 2). The main findings from analyses of Model 2 suggested (1) Learners performed better in the shared learning condition as compared to the individual learning condition (condition main effect:  $\gamma_{001} = 6.53$ ,  $t = 3.06$ ,  $df = 68$ ,  $p < .01$ ). Note, shared learning only occurred in the exploration phase (here run 1+2); the final two runs were completed individually by learners in all conditions. (2) *Distal learning effect*: All learners were able to improve

their performance from run 1+2 to run 4 indicating learning (simulation run main effect:  $\gamma_{010} = 2.32, t = 3.89, df = 130, p < .001$ ). (3) Whilst learners in the individual learning condition improved their performance from run 1+2 to run 4, they never reached the performance levels of the shared learning condition (condition x run interaction effect:  $\gamma_{011} = -2.49, t = -2.09, df = 130, p < .05$ ; simple effect at run 4  $F(1,68) = 5.31, p < .05$ , partial  $\eta^2 = .07$ , see Figure 2, left panel).

----

insert Figure 2 here

----

(4) *Proximal learning effect*: The advantage of the shared learning condition relative to the individual learning condition increased from trial to trial (condition x trial interaction effect:  $\gamma_{101} = 0.74, t = 2.95, df = 130, p < .01$ ). (5) Potential covariates (i.e., cognitive ability and personality) were also included in the model to explore whether they moderated the outcomes. No significant main or interaction effects of the covariates on performance were observed; the inclusion of these potential covariates in the model did not change any of the identified effects. For instance, when including abstract reasoning as a control variable the following results were obtained: main effect condition:  $\gamma_{001} = 6.72, t = 3.13, df = 67, p < .01$ ; main effect abstract reasoning:  $\gamma_{002} = -0.06, t = -0.93, df = 67, p = .35$ ; main effect simulation run:  $\gamma_{010} = 2.32, t = 3.89, df = 130, p < .001$ ; interaction effect condition X run:  $\gamma_{011} = -2.50, t = -2.09, df = 130, p < .05$ ; interaction effect condition X trial:  $\gamma_{101} = 0.74, t = 2.95, df = 130, p < .01$ ).

Finally (6), *Deliberation effect*: The analyses of Model 3 revealed that learners in the shared learning condition spent on average more time on each decision as

compared to learners in the individual learning condition (condition main effect:  $\gamma_{001}$   
 $= 8.35, t = 5.16, df = 68, p < .001$ ). Whilst this effect should not come as a surprise  
 where the exploration phase is concerned (due to the increased communication  
 demand in the shared learning condition) it still remains in run 3 (performance  
 phase) where everyone was working individually on the simulation (condition x run  
 interaction effect:  $\gamma_{011} = -5.12, t = -2.86, df = 68, p < .01$ ; simple effect at run 3:  $F(1,68) =$   
 $26.77, p < .001, \text{partial } \eta^2 = .28$ , see Figure 2, right panel). At run 4 both groups did not  
 differ significantly in their decision times  $F(1,68) = 0.50, p = .483, \text{partial } \eta^2 = .01$ ;  
 however learners who learned in the shared learning condition outperformed those  
 in the individual learning condition (see above), suggesting a higher level of  
 decision-making *efficiency* in the final run as a result of more effective learning.

In order to estimate the size of the shared learning effects on performance and  
 decision time, we compared the amount of explained variance in the final model  
 (Model 2, or Model 3 respectively) with the amount of explained variance in a model  
 that did not include the main and interaction effects of condition, but otherwise was  
 identical to Model 2 (or Model 3 respectively, i.e., a ‘baseline model’, see  
 proportional reduction in variance statistic, Peugh, 2010, p. 98). The reduction in  
 variance achieved by adding the main and interaction effects of condition to the  
 baseline models varied between 4.5%-5% for the performance model (Model 2) and  
 6.64%-32.02% for decision time model (Model 3, see footnote 2 for more detail).

In sum, findings suggest that learners learned to perform better in the  
 simulation when working in dyads whilst exploring the simulation. Although both  
 groups were instructed to use a hypothesis testing strategy, results support an  
 interpretation that learners in the shared learning group were more likely to employ

a hypothesis testing strategy than learners in the individual learning group. This is signified by the fact that learners who initially explored the simulation under the shared learning condition took longer for each individual decision in the first performance run compared to students who had explored the simulation individually. In this regard the superiority of their control performance can be seen as a result of increased reflectivity. In the second performance run, however, their reflectivity seems to turn into higher levels of efficiency as their decision-making time becomes shorter, reaching the level of their individually exploring counterparts, whilst their control performance remains superior.

#### 4 Discussion

Research suggests that unsystematic, ‘trial and error’ strategies lead to less effective learning which ultimately leads to inferior performance (Beckmann & Goode, 2014; de Jong & van Joolingen, 1998). However, even when explicitly instructed, learners do not always exploit the learning opportunities that more effective strategies afford. The aim of the current study was to design and test the effectiveness of a learning environment that encourages learners to exploit learning opportunities a training simulation is meant to provide. Wood et al. (2009) concluded from their analysis of the use of simulations in executive teaching, that simulations can be more effective when targeted facilitation is incorporated into the learning environment. We developed a dyadic learning setting that was structured, had minimal but targeted classroom facilitation (of the benefits of hypothesis testing strategies), and was inclusive of both learners. Performance outcomes under these conditions were compared to that of a control group who received identical

1 classroom facilitation, but worked through the simulations individually when first  
2 exploring the simulation. We were able to demonstrate that the shared learning  
3 approach tends to result in (a) initially longer decision times indicative of more  
4 systematic and deliberative information processing and decision-making, and  
5 consequently (b) better task performance.  
6  
7  
8  
9  
10

11       The use of simulations in higher education is not a novel idea. In fact, already  
12 15 plus years ago, virtually all business schools used simulations of some sort in their  
13 curricula (Faria, 1998; Faria, et al, 2009). Educators are typically very enthusiastic  
14 about the various benefits of using simulations in their teaching, and indeed,  
15 evidence suggests that students perceive simulations as highly motivating and  
16 engaging task environments (Cronan et al., 2011). Surprisingly, however, research  
17 evidence about the effectiveness of simulations in terms of measurable, objective  
18 learning outcomes is still limited to date (Bell et al., 2008). In addition to individual  
19 differences in learner attributes and the complexity of the task, it is the learning  
20 environment that has an impact on how learners approach the task and that  
21 subsequently determines the quality of their learning. Effective learning  
22 environments enable or even actively encourage learners to systematically process  
23 and utilise the information available in the simulation, including feedback.  
24  
25 Deliberative processing contributes to the effective acquisition of mental models that  
26 are necessary underpinnings of successful decision-making in complex  
27 environments. Hence, educators who wish to use simulations in their teaching need  
28 to ensure they design learning environments that are conducive to deliberative  
29 processing of information.  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65



The specific learning approach employed in the current study functions as a form of guidance in (a) how to *self-create* feedback by comparing explicitly predicted outcomes of decisions with actual outcomes, and (b) how to interpret potential discrepancies between the two as learning-relevant information. In other words, the rather quite simple approach of physically removing the decision-maker from the keyboard by implementing an intermediary in the interaction with the simulation improved the effectiveness of learning. We believe that the simplicity of this approach recommends its generic applicability across different types of simulations with the purpose of facilitating learning through increased deliberative processing of information (Romme, 2003; Wood et al., 2009). Future research on simulation-based learning should focus on (a) the replication of these findings using other simulations and more diverse samples, and (b) should explore the effectiveness of other approaches that prevent the adoption of a falsely understood “hands-on” focus and facilitates a more “brains-on” focus in learning with simulations.

Many real-world decision problems are complex (e.g., include many known and unknown variables) and dynamic in nature (i.e. their parameters change over time). Therefore, the ability to explicate and consciously deliberate on one’s decisions and their consequences, and to self-create on-going feedback on the quality of one’s decision-making under such conditions, is a capability that is highly relevant beyond the classroom. The current study demonstrates the effectiveness of a scripted interaction between members of a learning dyad that helped prevent the type of rapid decision-making that involves little *a priori*, systematic consideration of potential outcomes.

#### 4.1 Conclusion

The insight that hypothesis testing is advantageous when confronted with complex, dynamic problems is not new. However, we also know that hypothesis testing seems not to occur “naturally” when dealing with simulations, even when more or less explicitly taught. Likewise, simulations have been used in group settings before (e.g., Beckmann et al., 2012; Pasin & Giroux, 2011). However, simply encouraging students to work in dyads or groups is not enough. What makes learning in dyads potentially beneficial are settings in which a structured and systematic approach of communicating expectations, explicating assumptions, and justifying decisions is employed. The presented study provides an example for such an approach.

Research on the effectiveness of simulations as learning tools seems to revolve around the three questions – how simulations should be used, how simulations should be designed, and by whom simulations should be used. We believe increased efforts to address these questions and their interrelatedness is necessary. Adopting a more trans-disciplinary perspective that brings together expertise in psychology, system dynamics, and instructional design will be beneficial in facilitating effective learning with simulations in a wide range of educational contexts.

## 5 Footnotes

1. The Accelerated Learning Laboratory (A.L.L.) conducts expertise research and provides a 2-year leadership training program for mid-level managers from large organizations. The assessment and professional development component is a core feature of the program and has a theory based, elaborated assessment-for-learning focus. Part of the learning component of this program is the use of computer simulated micro-worlds in connection with various back-on-the-job activities where participants are encouraged to apply knowledge gleaned from lab experiences. The overarching objective was to foster the development of flexible expertise in managerial leadership that extends beyond domain-specific routine expertise.

2. Adding the main and interaction effects of condition to the performance baseline model led to roughly 5% reduction in both level-2 intercept variance ( $\tau(\pi)_{0jk\_baseline\_model2} = 114.31$ ,  $\tau(\pi)_{0jk\_full\_model2} = 108.49$ ) and level-2 slope variance ( $\tau(\pi)_{1jk\_baseline\_model2} = 2.53$ ,  $\tau(\pi)_{1jk\_full\_model2} = 2.38$ ), and 4.5% reduction in level-3 intercept variance ( $\tau(\beta)_{00\_baseline\_model2} = 27.12$ ,  $\tau(\beta)_{00\_full\_model2} = 25.89$ ). An insubstantial change in level-1 residual variance was observed ( $\sigma^2_{ijk\_baseline\_model2} = 116.95$ ,  $\sigma^2_{ijk\_full\_model2} = 116.96$ ). Similarly, adding the main and interaction effects of condition to the decision time baseline model led to 9.26% reduction in the level-2 intercept variance ( $\tau(\pi)_{0jk\_baseline\_model3} = 66.51$ ,  $\tau(\pi)_{0jk\_full\_model3} = 60.45$ ) and 6.64% reduction in the level-2 slope variance ( $\tau(\pi)_{1jk\_baseline\_model3} = 2.26$ ,  $\tau(\pi)_{1jk\_full\_model3} = 2.11$ ). Level-3 intercept variance decreased by 32.02% ( $\tau(\beta)_{00\_baseline\_model3} = 30.57$ ,  $\tau(\beta)_{00\_full\_model3} = 20.78$ ) and level-3 slope variance decreased by 19.63% ( $\tau(\beta)_{01\_baseline\_model3} = 33.52$ ,  $\tau(\beta)_{01\_full\_model3} = 26.94$ ). Again, no substantive

reduction in level-1 residual variance was observed ( $\sigma^2_{ijk\_baseline\_model3} = 429.16, \sigma^2_{ijk\_full\_model3} = 429.98$ ).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## 6 References

- Adobor, H., & Daneshfar, A. (2006). Management simulations: determining their effectiveness. *Journal of Management Development*, 25, 151-168.
- Anderson, J.R., & Schunn, C.D. (2000). *Implications of the ACT-R learning theory: no magic bullets*. In R. Glaser (Ed.), *Advances in instructional psychology*. Educational design and cognitive Science, Vol. 5, (pp. 1-34). Mahwah, NJ: Erlbaum.
- Beckmann, J.F., & Goode, N. (2014). The benefit of being naïve and knowing it: The unfavourable impact of perceived context familiarity on learning in complex problem solving tasks. *Instructional Science*, 42, 271-290.
- Beckmann, N., Wood, R. E., Minbashian, A., & Tabernero, C. (2012). Small group learning: Do group members' implicit theories of ability make a difference? *Learning and Individual Differences*, 22, 624-631.
- Bell, B. S., Kanar, A. M., & Kozlowski, S. W. J. (2008). Current issues and future directions in simulation-based training in North America. *International Journal of Human Resource Management*, 19, 1416-1434.
- Birney, D. P., Beckmann, J.F., & Wood, R. E. (2012). Precursors to the development of flexible expertise: Metacognitive self-evaluations as antecedences and consequences in adult learning. *Learning and Individual Differences*, 22, 563-574.
- Burns, B. D., & Vollmeyer, R. (2002). Goal specificity effects on hypothesis testing in problem solving. *The Quarterly Journal of Experimental Psychology*, , 55, 241-261.

- Chang, Y. C., Peng, H. Y., & Chao, H. C. (2010). Examining the effects of learning motivation and of course design in an instructional simulation game. *Interactive Learning Environments, 18*, 319-339.
- Chung, G., Harmon, T. C., & Baker, E. L. (2001). The impact of a simulation-based learning design project on student learning. *IEEE Transactions on Education, 44*, 390-398.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual*. Odessa, FL: Psychological Assessment Resources.
- Cronan, T. P., Douglas, D. E., Alnuaimi, O., & Schmidt, P. J. (2011). Decision-making in an integrated business process context: Learning using an ERP simulation game. *Decision Sciences Journal of Innovative Education, 9*, 227-234.
- De Jong, T., de Hoog, R., & de Vries, F. (1993). Coping with complex environments: The effects of providing overviews and a transparent interface of learning with a computer simulation. *International Journal of Man-Machine Studies, 39*, 621-639.
- De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*, 179-201.
- De Jong, T. (2006). Technological advances in inquiry learning. *Science, 312*, 532-533.
- Ellis, R. A., Marcus, G., & Taylor, R. (2005). Learning through inquiry: Student difficulties with online course-based material. *Journal of Computer Assisted Learning, 21*, 239-252.

- Fang, L., Tan, H. S., Thwin, M. M., Tan, K. C., & Koh, C. (2011). The value simulation-based learning added to machining technology in Singapore. *Educational Media International*, 48, 127-137.
- Faria, A. J. (1998). Business simulation games: Current usage levels - an update. *Simulation & Gaming*, 29, 295-308.
- Faria, A. J., Hutchinson, D., Wellington, W. J., & Gold, S. (2009). Developments in business gaming: A review of the past 40 years. *Simulation & Gaming*, 40, 464-487.
- Gresse van Wangenheim, C., Thiry, M., & Kochanski, D. (2009). Empirical evaluation of an educational game on software measurement. *Empirical Software Engineering* 14, 418-452.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2006). The international personality item pool and the future of public domain personality measures. *Journal of Research in Personality*, 40, 84-96.
- Gonzalez, C. (2005). Decision support for real-time, dynamic decision-making tasks. *Organizational Behavior and Human Decision Processes*, 96, 142-154.
- Gonzalez, C., Vanyukov, P., & Martin, M. K. (2005). The use of microworlds to study dynamic decision-making. *Computers in Human Behavior*, 21, 273-286.
- Goode, N., & Beckmann, J.F. (2010). You need to know: There IS a causal relationship between structural knowledge and control performance in complex problem solving tasks. *Intelligence*, 38, 345-352.
- Goodman, J., & Wood, R. E. (2004). Feedback specificity, learning opportunities, and learning. *Journal of Applied Psychology*, 89, 809-821.

- Goodman, J., Wood, R. E., & Hendrickx, M. (2004). Feedback specificity, exploration, and learning. *Journal of Applied Psychology*, 89, 248-262.
- Grössler, A. (2004). Don't let history repeat itself: Methodological issues concerning the use of simulators in teaching and experimentation. *System Dynamics Review*, 20, 263-274.
- Gureckis, T. M., & Markant, D. B. (2012). Self-Directed Learning: A cognitive and computational perspective. *Perspectives on Psychological Science*, 7, 464-481.
- Hung, W. (2008). Enhancing systems-thinking skills with modelling. *British Journal of Educational Technology*, 39, 1099-1120.
- Issenberg, S. B., Pringle, S., Harden, R. M., Khogali, S., & Gordon, M. S. (2003). Adoption and integration of simulation-based learning technologies into the curriculum of a UK Undergraduate Education Programme. *Medical Education*, 37(Suppl1), 42-49.
- Jansson, A. (1995). Strategies in dynamic decision-making: Does Teaching heuristic strategies by instructions affect performance? *Contributions to Decision Making*, 1, 213-232.
- Karakus, A., Duran, L., Yavuz, Y., Altintop, L., & Caliskan, F. (2014). Computer-based simulation training in emergency medicine designed in the light of malpractice cases. *Bmc Medical Education*, 14, 155.
- Lainema, T., & Nurmi, S. (2006). Applying an authentic, dynamic learning environment in real world business. *Computers & Education*, 47, 94-115.
- Lane, D. C. (1995). On a resurgence of management simulations and games. *Journal of the Operational Research Society*, 46, 604-625.



- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65
- Leemkuil, H., & de Jong, T. (2012). Adaptive advice in learning with a computer-based knowledge management simulation game. *Academy of Management Learning & Education*, 11, 653-665.
- Lerch, F. J., & Harter, D. E. (2001). Cognitive support for real-time dynamic decision-making. *Information Systems Research*, 12, 63-82.
- Leutner, D. (1993). Guided discovery learning with computer-based simulation games: Effects of adaptive and non-adaptive instructional support. *Learning and Instruction*, 3, 113-132.
- Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, 71, 449-521.
- McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. (2006). Effect of practice on standardised learning outcomes in simulation-based medical education. *Medical Education*, 40, 792-797.
- Mendonca, D., Chang, P., Hu, Y., & Gu, Q. (2012). Simulation, shared spaces, and information management to support discovery and learning: An application in emergency management. *Computer Applications in Engineering Education*, 20(2), 232-238.
- Njoo, M., & DeJong, T. (1993). Exploratory learning with a computer simulation for control theory: Learning processes and instructional support. *Journal of Research in Science Teaching*, 30, 821-844.
- Pasin, F., & Giroux, H. (2011). The impact of a simulation game on operations management education. *Computers & Education*, 57, 1240-1254.
- Paich, M., & Sterman, J. D. (1993). Boom, bust, and failures to learn in experimental markets. *Management Science*, 39, 1439-1458.

- Parush, A., Hamm, H., & Shtub, A. (2002). Learning histories in simulation-based teaching: The effects on self-learning and transfer. *Computers & Education*, 39, 319–332.
- Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, 48, 85-112.
- Pfahl, D., Klemm, M., & Ruhe, G. (2001). A CBT module with intergrated simulation component for software project management education and training. *Journal of Systems and Software*, 59, 283-298.
- Qudrat-Ullah, H. (2010). Perceptions of the effectiveness of system dynamics-based interactive learning environments: An empirical study. *Computers & Education*, 55, 1277-1286.
- Qudrat-Ullah, H., & Karakul, M. (2007). Decision-making in interactive learning environments: Towards an integrated model. *Journal of Decision Systems*, 16, 79-99.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.
- Ravert, P. (2002). An integrative review of computer-based simulation in the education process. *Computers, Informatics, Nursing*, 20, 203-208.
- Romme, A. G. L. (2003). Learning outcomes of microworlds for management education. *Management Learning*, 34, 51-61.
- Shellman, S. M., & Turan, K. (2006). Do simulations enhance student learning? An empirical evaluation of an IR simulation. *Journal of Political Science Education*, 2, 19-32.

- 1 Stouten, H., Heene, A., Gellynck, X., & Polet, H. (2012). Learning from playing with  
2 microworlds in policy making: An experimental evaluation in fisheries  
3 management. *Computers in Human Behavior*, 28, 757-770.  
4  
5  
6
- 7 Van Joolingen, W. R., & de Jong, T. (1993). Exploring a domain through a computer  
8 simulation: Traversing variable and relation space with the help of a  
9 hypothesis scratchpad. In D. Towne, T. de Jong & H. Spada (Eds.), *Simulation-  
10 based experiential learning* (pp. 191-206). Berlin: Springer.  
11  
12  
13  
14  
15  
16
- 17 Wastell, D., Peckover, S., White, S., Broadhurst, K., Hall, C., & Pithouse, A. (2011).  
18 Social work in the laboratory: Using microworlds for practice research. *British  
19 Journal of Social Work*, 41, 744-760.  
20  
21  
22  
23  
24
- 25 Wolfe, J. (1975). Effective performance behaviors in a simulated policy and decision-  
26 making environment. *Management Science*, 21, 872-882.  
27  
28  
29
- 30 Wood, R. E., & Bailey, T. C. (1985). Some unanswered questions about goal effects: A  
31 recommended change in research methods. *Australian Journal of Management*,  
32 10, 61-73.  
33  
34  
35  
36  
37
- 38 Wood, R. E., Beckmann, J. F., & Birney, D. P. (2009). Simulations, learning and real  
39 world capabilities. *Education+Training*, 51(5/ 6), 491-510.  
40  
41  
42
- 43 Zantow, K., Knowlton, D. S., & Sharp, D. C. (2005). More than fun and games:  
44 Reconsidering the virtues of strategic management simulations. *Academy of  
45 Management Learning & Education*, 4, 451-458.  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## 7 Figure Captions

Figure 1: Screenshot of the Employee Management Simulation (EMS)

Figure 2: Interaction effect condition X simulation run predicting performance (left panel) and decision time (right panel) in the simulation task. Note: Conditions only differ for the first data point, which represents the two run(s) in the exploration phase.

## Appendix A

The Level-1 equation for Model 2 was as follows (Model 3 differed in that the DV was decision time, not performance):

$$y_{ijk} = \pi_{0jk} + \pi_{1jk}(\text{trial}) + e_{ijk} \quad (1)$$

where  $y_{ijk}$  was performance in trial  $i$  in simulation run  $j$  of person  $k$ ,  $\pi_{0jk}$  was person  $k$ 's mean level of performance in run  $j$ ,  $\pi_{1jk}$  was the regression coefficient of trial on performance for person  $k$ 's run  $j$ , and  $e_{ijk}$  was an error term.

The Level-2 equations were as follows:

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}(\text{run}) + r_{0jk} \quad (2)$$

$$\pi_{1jk} = \beta_{10k} + r_{1jk} \quad (3)$$

where  $\beta_{00k}$  was the mean run performance of person  $k$ ,  $\beta_{01k}$  was the regression coefficient of run on performance,  $\beta_{10k}$  was the mean effect of trial on performance for person  $k$ , and  $r_{0jk}$  and  $r_{1jk}$  were random effects.

The Level-3 equations were as follows:

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{condition}) + u_{00} \quad (4)$$

$$\beta_{01k} = \gamma_{010} + \gamma_{011}(\text{condition}) + u_{01} \quad (5)$$

$$\beta_{10k} = \gamma_{100} + \gamma_{101}(\text{condition}) + u_{10} \quad (6)$$

where  $\gamma_{000}$  was the grand mean of performance across all simulation runs and all participants,  $\gamma_{001}$  was the main effect of condition,  $\gamma_{011}$  was the cross-level interaction between condition and run,  $\gamma_{101}$  was the cross-level interaction between condition and trial,  $u_{00}$ ,  $u_{01}$ ,  $u_{10}$  were random effects. Two error terms in Model 2 and one error term in Model 3 were fixed to zero ( $u_{01}$  and  $u_{10}$ , and  $u_{10}$ , respectively) because the associated variance was not significantly different from zero ( $\alpha = .01$ ).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## Acknowledgements

This research was supported under Australian Research Council's Linkage Projects funding scheme (project LP0669552). The views expressed herein are those of the authors and are not necessarily those of the Australian Research Council.

Table 1: Descriptive Statistics for Study Variables (*N* = 70)

Variables	<i>Mean</i>	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
1. Age	35.53	6.58														
2. Gender <sup>1</sup>	-	-	.33**													
3. Years of management experience	5.93	4.98	.63**	.24*												
4. Years in current job	1.99	1.66	-.06	.01	-.14											
5. Experimental condition <sup>2</sup>	-	-	.03	-	-.06	.08										
6. Verbal Reasoning	34.76	5.98	-.15	-.08	-.14	.03	-.01	(.82)								
7. Numerical Reasoning	16.80	5.84	-.38**	-.09	-.29*	.10	.01	.55**	(.91)							
8. Abstract Reasoning	23.36	6.40	-.29*	-.02	-.14	-.17	.14	.31**	.41**	(.85)						
9. Conscientiousness	74.75	11.59	-.03	.00	.04	.03	.10	-.24*	-.07	-.15	(.87)					
10. Neuroticism	30.51	13.72	-.05	-.03	-.12	-.02	-.09	.01	.13	.05	-.24 <sup>+</sup>	(.85)				
11. Openness	66.66	11.96	-.08	-.23 <sup>+</sup>	.16	-.04	-.11	-.14	-.13	.03	.14	-.28*	(.78)			
12. Agreeableness	70.65	10.38	.02	-.12	-.04	.07	-.13	-.16	-.09	-.15	.29*	-.45**	.17	(.76)		
13. Extraversion	64.29	15.55	-.31**	-.24*	-.05	-.11	.05	.18	.22 <sup>+</sup>	.30*	.07	-.28*	.20 <sup>+</sup>	.18	(.88)	
14. Simulation performance	54.03	9.72	.09	.07	.07	-.02	.41**	-.06	-.09	-.02	.18	-.13	-.04	-.11	.16	
15. Simulation decision time	22.39	7.95	.10	.03	-.03	-.11	.46**	.14	.04	-.09	.09	-.05	-.13	.04	-.07	.07

*Note:* <sup>1</sup>Gender was coded: 1=female (*N*=30), 2=male (*N*=40), <sup>2</sup>Experimental condition was coded 0=control, 1=experimental; for the variables gender and experimental condition point biserial correlations are reported; Coefficients in parenthesis represent Cronbach's  $\alpha$  based on larger samples (see Birney et al., 2012); \*\* $p < .01$ , \* $p < .05$ , <sup>+</sup> $p < .10$





Figure1

Figure 1

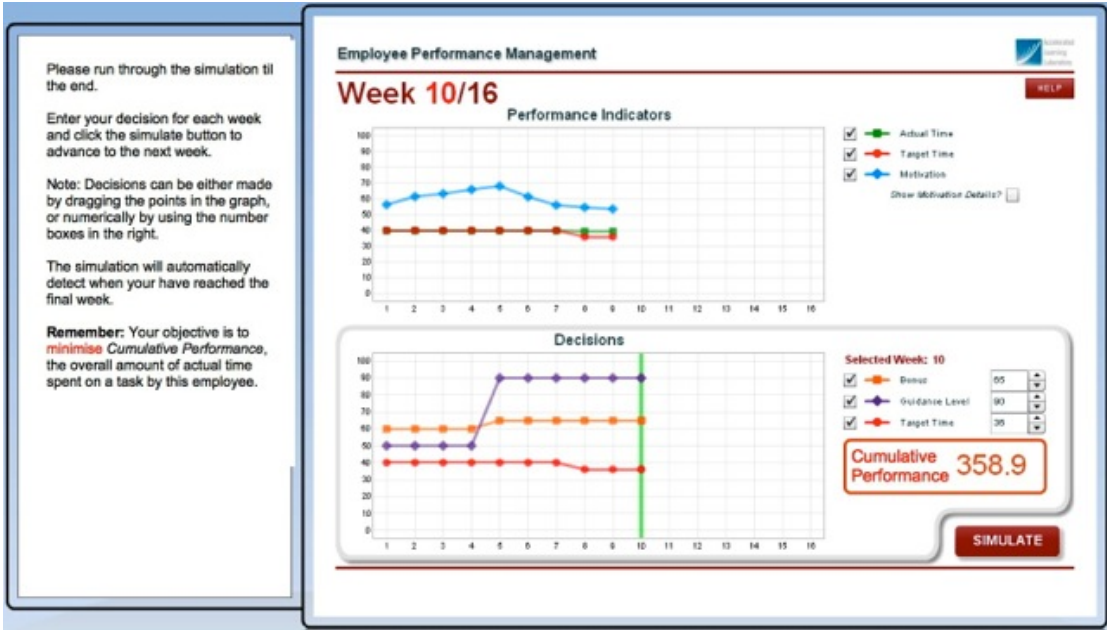
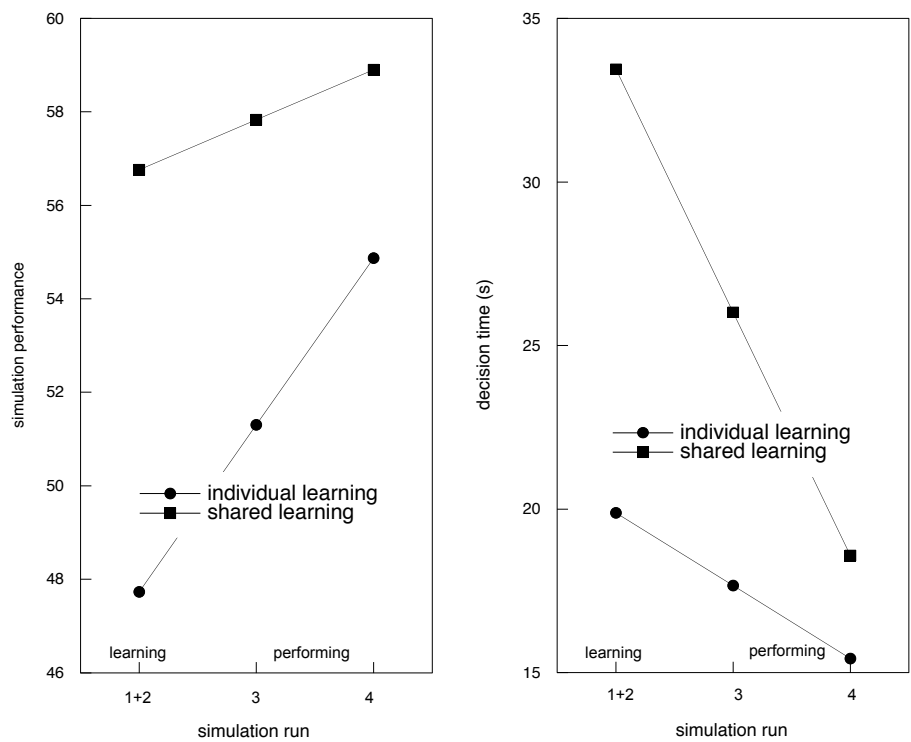


Figure2

Figure 2



## Highlights

- We address the issue of underutilisation of learning opportunities in simulations
- 71 professionals took part in an experiment using a management simulation
- Peer interactions were structured to encourage hypothesis-testing strategies
- Simple manipulation of how learners interact with the simulation affected learning
- Evidence for proximal, distal and deliberation learning effects is presented